DOCUMENT RESUME

ED 034 372 40 EC 004 786

AUTHOR MacKinnon, Ronald C.: Elliott, Charles
TITLE A Comparison of an Achievement Battery with Two Tests of Ability with Educable Mental Retardates. Final Report.
INSTITUTION Florida State Univ., Tallahassee.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Bureau of Education for the Handicapped.
PUB DATE Aug 69
GRANT OEG-0-9-190611-0774(032)
NOTE 14p.

EDRS PRICE EDRS Price MF-$0.25 HC-$0.80
DESCRIPTORS Achievement Tests, *Aptitude Tests, Comparative Testing, Correlation, Educable Mentally Handicapped, *Exceptional Child Research, Group Tests, Individual Tests, Intelligence Tests, *Mentally Handicapped, Testing, Testing Problems, *Test Validity

ABSTRACT

To find the concurrent validity of two scholastic aptitude tests when a scholastic achievement test was used as a criterion for use in placement of mentally retarded children, 127 subjects were involved. The California Achievement Test (CAT) was used as a criterion measure, and the Primary Mental Abilities test (PMA) and the Slosson Intelligence Test (SIT) were used to determine the correlation of the scores with the criterion measure. The results indicated correlations of .68 between the PMA and the CAT and .62 between the SIT and the CAT. The intercorrelation of the PMA and SIT was .67. Ability subtests were analyzed to determine their correlations to the criterion. The conclusion was that when six or more children are to be tested, less time if needed and no concurrent validity is lost by using the PMA rather than the individual test, the SIT. Tables of results are included. (Author/JM)

PA-40

ED034372

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

FINAL REPORT

Project No. 19-2061
Grant No. OEG-0-9-192061-0774(032)

A COMPARISON OF AN ACHIEVEMENT BATTERY WITH TWO TESTS OF ABILITY WITH EDUCABLE MENTAL RETARDATES

Ronald C. MacKinnon and Charles Elliott
Florida State University
Tallahassee, Florida 32306

August 1969

Department of Health, Education, and Welfare

U.S. Office of Education
Bureau of Education for the Handicapped

ED 034 372

ED034372

19-2061
PA 40
OE/BEH

Final Report

Project No. 19-2061
Grant No. OEG-0-9-192061-0774(032)

A COMPARISON OF AN ACHIEVEMENT BATTERY WITH TWO TESTS OF ABILITY WITH EDUCABLE MENTAL RETARDATES

Ronald C. MacKinnon and Charles Elliott
Florida State University
Tallahassee, Florida 32306

August 1969

The research reported herein was performed pursuant to a Grant with the Bureau of Education for the Handicapped, U.S. Office of Education, Department of Health, Education and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official position of the Bureau of Education for the Handicapped.

Department of Health, Education and Welfare

U.S. Office of Education
Bureau of Education for the Handicapped

CONTENTS

Subject	Page
SUMMARY	
INTRODUCTION	1
METHOD	2
ANALYSIS AND FINDINGS	3
CONCLUSIONS AND RECOMMENDATIONS	4
REFERENCES	10

LIST OF TABLES

TABLE 1 Means and Standard Deviation of all Subtests and Totals 5

TABLE 2 Intercorrelation Matrix of Predictor Subtests and Totals with Criterion 6

TABLE 3 Summary Table of the Stepwise Multiple Regression 7

TABLE 4 Results of the Multiple Regression Analysis with CAT Subtests as Criteria 8

A COMPARISON OF AN ACHIEVEMENT BATTERY WITH TWO TESTS OF ABILITY WITH EDUCABLE MENTAL RETARDATES

Summary

The purpose of this study was to find the concurrent validity of two scholastic aptitude tests when a scholastic achievement test was used as a criterion and when the tests were applied to children with mental retardation.

The use of the instrument with the greatest concurrent validity will increase the efficiency and accuracy with which mentally retarded children can be placed in special classes.

Children (184) in sixteen elementary special classes in Leon County, Florida were administered a criterion measure, the California Achievement Test. The Primary Mental Abilities test and the Slosson Intelligence Test were also administered to determine the correlation of the scores with the criterion measure, and intercorrelations between tests.

Attrition, due to absenteeism and an inability of the subjects to understand the verbal directions of the CAT reduced the number of subjects to 127.

Analysis of the data using a stepwise multiple regression computer program indicated that the PMA correlated r=0.676 and the SIT correlated r=0.618 with the criterion measures. These correlations were not significantly different at the .05 level of confidence. The intercorrelation of the PMA and SIT was r=0.665. A secondary analysis determined which ability subtest or total correlated most highly with each of the nine subscores of the criterion variable. The means and standard deviations of all subscores were found.

It was concluded that when six or more children are to be tested an economy of test administration time may be effected without loss of concurrent validity, by employing the group test (PMA) rather than the individual test (SIT).

A COMPARISON OF AN ACHIEVEMENT BATTERY WITH TWO TESTS OF ABILITY WITH EDUCABLE MENTAL RETARDATES

Introduction

According to Heber's definition, to be classified as mentally retarded, a child must display subnormal intellectual functioning, which is usually measured by a standardized intelligence test, and maladaptive behavior of which school achievement is a primary criterion in school age children. Customarily school achievement is measured by using a standardized achievement test. The intelligence test was devised as a predictor of school achievement and must therefore be classified as a test of scholastic aptitude rather than a test of intelligence. Consequently in order to classify a child as mentally retarded, an aptitude and a post-facto test of school success must coincide. School achievement tests and scholastic aptitude tests should, therefore, correlate highly with one another when they are applied to children with mental retardation.

The purpose of this study was to find the concurrent validity of two scholastic aptitude tests and a scholastic achievement test when the tests were applied to children with mental retardation. The intelligence test having the higher correlation with an achievement test would be the better aptitude test for identifying such children.

Mueller (1965) investigated the empirical validity of six psychometric tests, as predictors of learning ability: The Stanford Binet Intelligence Scale (S-B, Terman & Merrill, 1960); the Illinois Test of Psycholinguistic Abilities (ITPA, McCarthy & Kirk, 1961); the Peabody Picture Vocabulary Test (PPVT, Dunn, 1959); the Pictorial Test of Intelligence (PTI, French, 1963); the Colored Progressive Matricies (CPM, Raven, 1947); and the Primary Mental Abilities Test (PMA, Thurstone and Thurstone, 1962). He found, "The picture which emerged in these comparisons suggests that the PMAT, S-B, PTI, and ITPA are the most valid of the predictor tests under consideration in predicting learning ability in young mental retardates. . . . the PMAT was superior to both the Binet and PPVT in predicting reading and arithmetic achievement." (Mueller, 1965).

The PMA (Thurstone & Thurstone, 1962) is a teacher-administered group screening instrument of mental maturity. In view of the correlation (0.79) reported by Mueller for the PMA with his learning criteria, and reading and arithmetic achievement, investigation of the concurrent validity of the PMA when compared with other teacher-administered instruments especially the more time consuming individual tests, should be conducted.

One such screening instrument which seems to be gaining considerable prominence as a teacher-administered test, is the Slosson Intelligence Test (SIT, Slosson, 1963). Admission to special education classes in some states is determined on the basis of an individual psychological examination by a qualified examiner. The SIT meets this requirement in that it is an individually administered psychological test. The manual lists teachers, social workers, guidance counselors and others as qualified examiners. However, the SIT has been used on the basis of its face validity alone; e.g., there is currently no statement of the SIT validity or statistical reliability in the literature (Baumeister, 1967). To be valid for the purpose of admitting to special classes, it is necessary that the SIT, as does the PMA, correlate highly with a scholastic achievement test. In this study the PMA form K-1 was found to be appropriate for the subjects. The SIT is obtainable in only one form.

An achievement test which has been used extensively and which lends itself well to educational planning, due to the diagnostic analysis page included in the test protocol, is the California Achievement Test (CAT, Tiegs & Clark, 1957 with 1963 norms). The CAT has reported test retest reliability coefficients of from .97 to .98 on the three lower ranges of the battery. The lower primary Form W, grades one and two, was used.

The scores obtained on the CAT measure served as an accepted contemporary criterion of performance. The concurrent validity of the SIT was determined by correlating the scores obtained on the SIT with those obtained on the criterion measure. The concurrent validity of the PMA obtained in a like fashion. The intercorrelations between the PMA and the SIT scores were also determined.

Usually a school teacher cannot administer an S-B test in a valid manner, which is unfortunate because teachers make the greatest use of the intelligence quotient and it is in the S-B that psychometrists usually have the most faith. Mueller's study may be seen as part of a real, though unorganized and perhaps not always conscious, nationwide program to determine which teacher administered IQ test is most deserving of the teacher's faith. The search is based on the logic that if children benefit from teacher-gathered IQ scores, then the benefit should be a positive function of the validity and reliability of the tests employed. Mueller's study indicates that the PMA has high concurrent validity with his criteria.

The present study extends Mueller's work by comparing the concurrent validity, using a more commonly found criterion, of the SIT and the PMA. The SIT can be characterized as a brief S-B type of teacher administered test.

Method

The experimental subjects consisted of the total number (184) of children who were enrolled in the sixteen elementary public school

special classes for the educable mentally retarded in Leon County, Florida, all experimental testing was accomplished in two weeks per class. Those subjects not present for all testing sessions were excluded from the sample. To control for systematic bias due to non-random absenteeism, no testing was done during weeks that included national or religious holidays which would have tended to systematically exlude a segment of the sample. Attrition was high due to absenteeism and a seeming inability for some children to understand the verbal instructions necessary for the administration of the CAT.

Administration of the test instruments was accomplished by trained student examiners. All of the tests were administered to all of the subjects. The group tests were administered to each class in the manner described in the manuals. The SIT was individually administered in the standard manner. The order of test administration was randomized. The protocols were scored and recorded by the examiners and then re-scored by the investigator to assure accuracy and consistency.

Complete data was collected from 146 subjects. Inspection of the protocols by an expert revealed that some protocols had to be judged invalid. This judgment was usually made on the basis that the child must have misunderstood the task. When all protocols were checked, complete and valid data were available for 127 subjects. These were subjected to statistical analysis.

Analysis and Findings

All scores were derived as age equivalents from norms published in their respective manuals. The data from all tests and subtests, where necessary, were converted to age equivalents in months by the rule of five (Dunn, 1964, p.60). Thus the third month of grade one yielded an achievement age of 6 years 3 months or 75 months. These data were analysed in a large computer using Efroymson's program for stepwise multiple regression. Table 1 displays the means and standard deviations for subtests and totals of all three tests. The table indicates that a ratio IQ obtained from the formula $\frac{MA}{CA} \times 100 = IQ$, yields a mean IQ of 71 for the total population when Slosson MA is used, and 65 IQ when PMA scores are used. On only one collection of data, the Perceptual Speed subtest of the PMA, did the subjects perform better that would be anticipated by their SIT score and its grade equivalent by the rule of five. On all other subtests and total scores, the performance was below this expectation. In this study the total score of the PMA was a more accurate predictor of school achievement than was the SIT even though both tended to overestimate the achievement. The total battery mean score for the CAT was 78 months, the PMA total was 83 months and the SIT total was 91 months. These data indicate that in this particular study more accurate predictions of CAT scores could have been obtained by using the PMA in a group testing situation than by using the individually administered SIT. However, these results can not be generalized beyond this study because the correlations were not significantly different at the

.05 level.

Table 2 shows the correlations between each ability subtest and total when the CAT total is used as the criterion variable. The table indicates that the PMA total score correlates r=0.676 and the SIT total correlates r=0.618, with the criterion. A test of the differences of correlations by the method of Walker and Lev (1958) revealed that the apparently higher concurrent validity of the PMA could have this appearance by chance alone (P=<.10). The most that can be said from this study is that the concurrent validity of the two tests of intelligence do not differ by more than can be attributed to the chance factors inherent in the tests, the administrations of the tests, and other experimental procedures. It does appear to be certain, however, that if more than six children are to be examined, an economy of administration time with no sacrifice of validity can be obtained through the use of the group PMA.

Table 3 shows the partitioning of the total sum of squares into portions attributable to each of the contributing factors. It verifies that in the present study the PMA total score was the best predictor of the criterion variable, and that adding knowledge of the SIT score yielded an increase in r square of only .05. The intercorrelation of the PMA and SIT is r=0.665 (Table 2). It must therefore be concluded that the practice of non-selectively administering both tests is uneconomical of time in that knowledge of the SIT score adds but little increased confidence in obtaining a valid measure, when the PMA score is known.

A secondary concern of this study was to determine which ability subtest or total, correlated most highly with each of the nine subscores of the criterion variable. These results are presented in Table 4 which includes only those factors that increase r square by more than .01. The table indicates that the SIT total score was not the best predictor of any of the subscores or the total of the CAT. The PMA total score was the single best predictor of Reading Vocabulary subtest and the total battery. The PMA Verbal Meaning subtest was the single best predictor of Reading Comprehension, Total Reading, Mechanics of English, Spelling, and Total Language subscores of the CAT. The PMA Number Facility subtest was the single best predictor of Arithmetic Reasoning, Arithmetic Fundamentals and Arithmetic Total, subscores of the CAT. The SIT total score was the second best indicator of these same arithmetic subscores.

Conclusions and Recommendations

The experimenter is not satisfied with the results of this study. The testing using the SIT progressed easily and satisfactorily as did the testing with the PMA, but with the Southern elementary EMR children, 60 per cent of whom were Negro, administering the CAT, which was the important criterion measure, was fraught with difficulties. The youngest one-quarter of the population seemed to be working on the ragged edge of their abilities when simply comprehending the verbal instructions. Attrition selec-

Table 4

Means and Standard Deviation of all Subtests and Totals

Variable	Mean	Standard Deviation
PMA		
Verbal Meaning	80.94	14.04
Perceptual Speed	93.02	16.91
Number Facility	83.91	14.72
Spatial Relations	80.19	20.78
Total Score	83.20	14.85
SIT		
MA	90.77	15.26
CA	127.42	21.40
CAT		
Total Reading Vocabulary	78.33	7.22
Reading Comprehension	74.26	8.29
Total Reading	77.83	6.55
Arithmetic Reasoning	79.06	6.62
Arithmetic Fundamentals	82.31	12.03
Total Arithmetic	79.54	8.08
Total Mechanics of English	77.46	6.77
Spelling	76.51	12.23
Total Language	76.98	7.49
Total Battery	77.87	6.84

Table 2

Intercorrelation Matrix of Predictor Subtests and Totals with Criterion

VARIABLE	Verbal Meaning	Perceptual Speed	Number Facility	Spatial Relations	PMA Total	CAT Total	SIT Total	CA Total
NUMBER	1	2	3	4	5	6	7	8
1	1.000	.491	.668	.685	.801	.659	.590	.497
2		1.000	.622	.654	.754	.499	.425	.191
3			1.000	.670	.819	.650	.652	.279
4				1.000	.897	.635	.551	.375
5					1.000	.676	.665	.369
6						1.000	.618	.321
7							1.000	.351
8								1.000

Table 3

Summary Table of the Stepwise Multiple Regression

Step Number	Variable Entered	Removed	Multiple R	Increase In RSQ	F Value To Enter or Remove
1	PMA Total		.6764	.4576	105.45
2	SIT		.7129	.0506	12.77
3	PMA Verbal Meaning		.7324	.0282	7.48
4	PMA Spatial Relations		.7414	.0133	3.60
5	PMA Total		.7414	-.0000	.00
6	PMA Number Facility		.7516	.0152	4.28
7	PMA Total		.7554	.0057	1.60
8	PMA Perceptual Speed		.7596	.0063	1.78
9	CA		.7602	.0009	.26

Table 4

Results of the Multiple Regression Analysis, with CAT Subtests as Criteria

Step Number	Variable Entered	Multiple R	Increase In RSQ	F Value To Enter
Criterion:	Reading Vocabulary			
1	PMA Total	.6248	.3904	80.0417
2	Verbal Meaning	.6491	.0310	6.6418
3	Spatial Relations	.6693	.0266	5.9291
4	SIT MA	.6801	.0145	3.2942
Criterion:	Reading Comprehension			
1	Verbal Meaning	.6233	.3885	79.4093
2	Spatial Relations	.6480	.0314	6.7216
Criterion:	Total Reading			
1	Verbal Meaning	.6300	.3969	82.2764
2	Spatial Relations	.6629	.0694	16.1212
3	SIT MA	.6947	.0163	3.8677
Criterion:	Arithmetic Reasoning			
1	Number Facility	.7265	.5278	139.7382
2	SIT MA	.7665	.0907	29.4973
3	Spatial Relations	.8052	.0298	10.4081
Criterion:	Arithmetic Fundamentals			
1	Number Facility	.7297	.5324	142.3181
2	SIT MA	.7557	.0387	11.2028
3	PMA Total Score	.7663	.0161	4.7947
Criterion:	Total Arithmetic			
1	Number Facility	.7760	.6022	189.2239
2	SIT MA	.8223	.0740	28.3288
3	Spatial Relations	.8347	.0205	8.3161
Criterion:	Mechanics of English			
1	Verbal Meaning	.6164	.3799	76.5887
2	SIT MA	.6559	.0503	10.9507
3	Chronological Age	.6819	.0348	8.0044
4	Perceptual Speed	.7020	.0277	6.6644
Criterion:	Spelling			
1	Verbal Meaning	.5077	.2578	43.4175
2	Number Facility	.5265	.0195	3.3381
Criterion:	Total Language			
1	Verbal Meaning	.6064	.3677	72.6895
2	Number Facility	.6417	.0440	9.2791
3	SIT MA	.6540	.0160	3.4365
4	Chronological Age	.6636	.0130	2.8272

tively biased the population. It weeded out the less verbal children with MA's below seven years and those children with the strongest local dialect. These observations, made by the principal investigator rather than the actual data itself, has lead the experimenter to conclude that the CAT was a poor choice of criterion measure for such a population. Too many children were working at the bottom of the test apparently because of the nature of the instructions and the nature of their language behavior. Whenever protocols were suspected of being contaminated by this influence, they were discarded. It is suspected that the CAT total score would have been higher if the above language problem had not been encountered, in which case the obtained correlations would have altered somewhat.

Speculation based on the observable trends leads to the supposition that the alteration might have been in the direction of Mueller's findings of r=0.79.

In spite of the attrition problem it is safe to conclude that when six or more children are to be examined, an economy of test administration time can be gained without loss of validity by using the group test rather than the individual test. Those trends that exist all indicate that the PMA may be more valid than the SIT and is not less valid, when used with Southern, elementary level, EMR children.

REFERENCES

Baumeister, A. A. Mental Retardation. Appraisal, Education and Rehabilitation. Chicago: Aldine Publishing Company, 1967.

Binet, A. and Simon, T. La necessite d'etablir un diagnostic scientifique sur les etats inferieurs d'intelligence. L'Annee Psychologigue, 1905, 163-191.

Dunn, L, M. Peabody Picture Vocabulary Test. American Guidance Service, Inc., 1959.

Dunn, L. M. Exceptional Children in the Schools. New York: Holt-Rinehart-Winston, 1963.

French, J. L. Pictorial Test of Intelligence. Houghton Mifflin Co., 1963.

McCarthy, J. J. and Kirk, S. A. Illinois Test of Psycholinguistics Abilities, Experimental Edition. University of Illinois Press, 1961.

Mueller, M. W. A comparison of the empirical validity of six tests of ability with young educable retardates. IMRID Behavioral Science Monograph, 1965, No. 1.

Raven, J. C. Coloured Progressive Matrices. Psychological Corporation, 1947.

Slosson, R. L. Slosson Intelligence Test for Children and Adults. Slosson Educational Publications, 1963.

Terman, L. M. and Merrill, Maud A. Stanford-Binet Intelligence Scale, Houghton Mifflin Company, 1960.

Tiegs, E. W. and Clark, W. W. California Achievement Tests. California Test Bureau, 1963.

Thurstone, L. L. and Thurstone, Thelma G. Primary Mental Abilities, Revised. Science Research Associates, Inc., 1962.

Walker, H. M. and Lev, J. Elementary Statistical Methods. (Revised Ed.) New York: Holt-Rinehart-Winston, 1958.